การสรุปใจความสำคัญของการ์ตูนเคลื่อนไหวด้วยขนาดที่หลากหลายของคีย์เฟรม

นายภาคย์ภูมิ ตาณพิเชษฐ์

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิทยาศาสตรมหาบัณฑิต
สาขาวิชาวิทยาการคอมพิวเตอร์และเทคโนโลยีสารสนเทศ
ภาควิชาคณิตศาสตร์และวิทยาการคอมพิวเตอร์
คณะวิทยาศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย
ปีการศึกษา 2555

CARTOON ANIMATION SUMMARIZATION USING MULTI-SCALED KEYFRAMES

Mr. Pakpoom Tanapichet

A Thesis Submitted in Partial Fulfillment of the Requirements
for the Degree of Master of Science Program in
Computer Science and Information Technology
Department of Mathematics and Computer Science
Faculty of Science
Chulalongkorn University
Academic Year 2012

Thesis Title	CARTOON ANIMATION SUMMARIZATION USING MULTI-SCALED KEYFRAMES
By	Mr. Pakpoom Tanapichet
Field of Study	Computer Science and Information Technology
Thesis Advisor	Assistant Professor Rajalida Lipikorn, Ph.D.
Thesis Co-advisor (if any)	Assistant Professor Nagul Cooharojananone, Ph.D.

Accepted by the Faculty of Science, Chulalongkorn University in Partial Fulfillment of the Requirements for the Master's Degree

………………………………………….. Dean of the Faculty of Science
(Professor Supot Hannongbua, Dr.rer.nat.)

THESIS COMMITTEE

………………………………………….. Chairman
(Suphakant Phimoltares, Ph.D.)

………………………………………….. Thesis Advisor
(Assistant Professor Rajalida Lipikorn, Ph.D.)

………………………………………….. Thesis Co-advisor
(Assistant Professor Nagul Cooharojananone, Ph.D.)

………………………………………….. External Examiner
(Assistant Professor Panjai Tantasanawong, Ph.D.)

ภาคย์ภูมิ ตาณพิเชษฐ์ : การสรุปใจความสำคัญของการ์ตูนเคลื่อนไหวด้วยขนาดที่หลากหลายของคีย์เฟรม. (CARTOON ANIMATION SUMMARIZATION USING MULTI-SCALED KEYFRAMES) อ. ที่ปรึกษาวิทยานิพนธ์หลัก : ผศ. ดร. รัชลิดา ลิปิกรณ์ , อ. ที่ปรึกษาวิทยานิพนธ์ร่วม : ผศ. ดร. นกุล คูหะโรจนานนท์, 48 หน้า.

งานวิจัยนี้ได้ทำการนำเสนอวิธีการสร้างคีย์เฟรมแบบใหม่ โดยมีวัตถุประสงค์ในการเพิ่มรายละเอียดของเนื้อหาในคีย์เฟรมให้มีความสมบูรณ์มากยิ่งขึ้น ซึ่งการสรุปใจความสำคัญของวีดีทัศน์ทั่วไปนั้น มักจะมีข้อมูลภาพตกหล่นด้วยข้อจำกัดทางอัตราส่วนของวีดีทัศน์ จึงเป็นที่มาของการปรับปรุงให้ดียิ่งขึ้น โดยใช้การนำเสนอด้วยภาพปริทัศน์ซึ่งบรรจุข้อมูลได้มากกว่าภาพปกติ แนวคิดหลักคือการทำสัญลักษณ์บนรหัสเวลาของวีดีทัศน์ด้วยข้อมูลรอยต่อระหว่างฉากและทิศทางการเคลื่อนไหวของวัตถุ ช่วงเวลารอยต่อระหว่างแต่ละสัญลักษณ์ที่ถูกทำไว้จะถูกปฏิบัติเสมือนว่าเป็นหนึ่งฉาก โดยทำงานร่วมกับข้อมูลการเคลื่อนไหวของวัตถุเพื่อคัดเลือกเฟรมสำคัญในการสร้างคีย์เฟรมและคีย์เฟรมปริทัศน์ แต่ละคีย์เฟรมที่ถูกสร้างขึ้นจะถูกใช้ในรูปแบบของช่องหนังสือการ์ตูน และช่องหนังสือการ์ตูนเหล่านี้จะถูกรวบรวมและเรียบเรียงเพื่อการนำเสนอในรูปแบบหนังสือการ์ตูน ระบบที่นำเสนอในงานวิจัยนี้ใช้ลักษณะการเรียบเรียงที่หลากหลาย โดยมีเป้าหมายในการใช้พื้นที่ของหนังสือการ์ตูนที่ถูกสร้างขึ้นมาอย่างเหมาะสม ผลสุดท้ายที่ได้รับจากงานวิจัยนี้คือหน้ากระดาษของหนังสือการ์ตูนที่บรรจุเนื้อหาจากการ์ตูนเคลื่อนไหวที่สามารถบรรยายได้ในลักษณะเดียวกับหนังสือการ์ตูน

ภาควิชาคณิตศาสตร์และวิทยาการคอมพิวเตอร์ ลายมือชื่อนิสิต...

สาขาวิชาวิทยาการคอมพิวเตอร์ ลายมือชื่อ อ.ที่ปรึกษาวิทยานิพนธ์หลัก...........

และเทคโนโลยีสารสนเทศ ลายมือชื่อ อ.ที่ปรึกษาวิทยานิพนธ์ร่วม...........

ปีการศึกษา 2555

5273627123 : MAJOR COMPUTER SCIENCE AND INFORMATION TECHNOLOGY
KEYWORDS : VIDEO SUMMARIZATION / CARTOON ANIMATION / KEYFRAME GENERATION / PANORAMA IMAGE / SHOT BOUNDARY / OPTICAL FLOW

PAKPOOM TANAPICHET : CARTOON ANIMATION SUMMARIZATION USING MULTI-SCALED KEYFRAMES. ADVISOR : ASST. PROF. RAJALIDA LIPIKORN, Ph.D., CO-ADVISOR : ASST. PROF. NAGUL COOHAROJANANONE, Ph.D., 48 pp.

This research proposes a novel method to generate keyframes from cartoon animation with the aim to improve the details and completeness of contents represented by keyframes. Consider that general techniques on video summarization usually discard some important contents due to its restriction on aspect ratio; this research thus proposes a new method using panorama technology to add more details to be included in each keyframe. The concept is to mark the time code based on shot boundary and optical flow direction. The period of time between every two consecutive marked time codes is used to form a shot sequence which is actually a sequence of frames. The global and local optical flows are also used to determine how to select the frames and when to stitch the frames together according to the rules. Each of these generated keyframes is treated as a comic panel and they are organized to be represented in comic book style. The proposed algorithm aims to fit several comic panels into a comic page while also optimizing its spaces using various types of arrangements. The results of this proposed method are the pages consist of comic panels that aesthetically represent the cartoon animation in a similar way to their respective comic adaptations.

Department : Mathematics and Computer Science
Field of Study : Computer Science and Information Technology
Academic Year : 2012

Student's Signature ....
Advisor's Signature ....
Co-advisor's Signature ....

ACKNOWLEDGEMENT

I would like to use this space to express my utmost appreciation to my thesis advisor, Assistant Professor Rajalida Lipikorn and thesis co-advisor Assistant Professor Nagul Cooharojananone for their guidance and support on this research the whole time at Department of Mathematics and Computer Science. In addition, I would also like to thank the thesis committees, Dr. Suphakant Phimoltares and Assistant Professor Panjai Tantasanawong for taking their precious time in thesis examinations and for their invaluable comments.

I would also like to thank all reviewers of the IEEE-ISPACS 2011 and the IS&T/SPIE 2012 Electronic Imaging conferences for their constructive comments on the published proceedings that guide some improvements on this research. Finally, I would like to show my gratitude to TKO Comics Co., Ltd. and TIGA Co., Ltd. for their suggestions regarding the usage of copyrighted materials and for permission of image usages in this research.

CONTENTS

Chapter Page

LIST OF TABLES

LIST OF FIGURES

Chapter 1

INTRODUCTION

1.1 Background and Importance

Digital videos are being widely consumed among computer users as one of the most popular media as they contain information in finer details and are also convenient to carry. A video generally contains full cover of information users need; however, it is very time consuming to search for a specific content from a video. In order to use video as an important source of information, researches on digital videos have become interesting topics among researchers and video summarization is one of them. Several techniques on video summarization have been introduced in the literatures. Most of the existing techniques [1, 2, 3, 4, 5, 6, 7] involved with video summarization tend to select the most suitable frames (also known as “keyframes”) from a video and use them to represent the whole video. These techniques restrict themselves to the video aspect ratio. The keyframes are usually selected based on how important the information contained in that specific frame is and how the researchers define it. These techniques select the keyframes directly from a shot sequence where they maintain their dimension according to the video aspect ratio.

1.2 Problem Formulation

The problem arises when a shot sequence is spanning for a long period of time then some important elements might be missing from a keyframe if only one keyframe is selected. On the other hand, if two keyframes are selected then the elements that are supposed to be on the same keyframe are divided into two keyframes due to the aspect ratio, this can make the content in each keyframe to be misinterpreted and the content in each keyframe will not make much sense as shown in Figure 1.1(a)-(b).

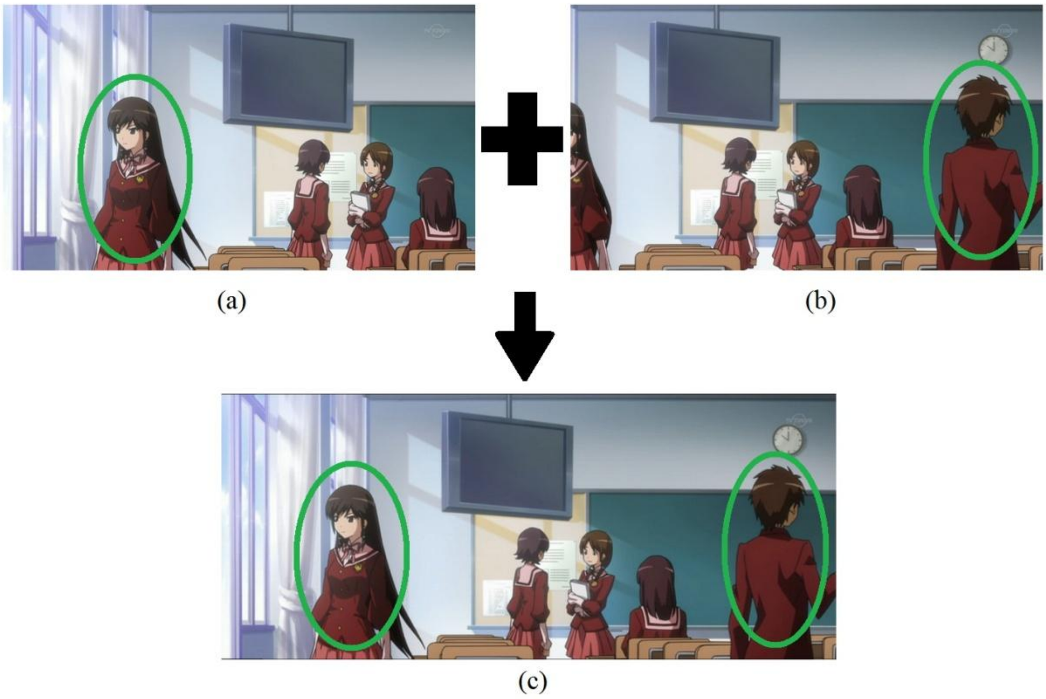

Figure 1.1: Comparison between (a) and (b) fixed-aspect-ratio keyframes and (c) the panorama keyframe with two important elements circled. ©Tamiki Wakaki/Shogakukan, Kaminomi project, TV TOKYO

The existing techniques mentioned above cannot represent panning shots sensibly. Some techniques divide important elements that are supposed to be together into separated keyframes. Some techniques omit those important elements for the more important one.

1.3 Proposed Solution

The main idea of this research is to put those elements together in one keyframes when they are supposed to. It proposes a new technique to generate keyframes from cartoon animation that can cope with multiple important contents. The proposed technique does not select keyframes from a collection of frames, but generate new keyframes according to the rule-based optical flow. Moreover, the proposed method can generate panorama keyframes when a camera is detected as panning. Each of panorama keyframes combines multiple important elements into one keyframe as shown in Figure 1.1(c). As a result, the keyframes representing one video will have various sizes depending on optical flows, and the sizes of panorama keyframes are supposed to be larger.

Figure 1.1 shows how the proposed technique generates a panorama keyframe that includes every important element from each frame. If a frame is restricted by the aspect ratio then it is obviously unable to contain both important elements. This problem can be eliminated by using a panorama keyframe which can depict continuity of information better than multiple non-panorama keyframes.

1.4 Objectives

1. To improve cartoon animation summarization to cover more contents within each keyframe.

2. To generate multi-scaled keyframes that can cover more contents in the panning shots.

1.5 Scope of the Work

1. The samples are restricted to 2D cartoon animation.

2. The samples are restricted to Japanese cartoon animation for a reason of material availability. There is no other type of cartoon animation available which exists a direct comic adaptation to act as a comparable sample.

3. The object motions are excluded from panning sequence; i.e. object cannot move while the shot is panning as discussed in chapter 4.

4. The panning shot cannot revolve around object as discussed in chapter 4.

5. Audio feature is excluded.

1.6 Expected Outcomes

1. Multi-scaled keyframes which have various sizes.
2. Better representation of cartoon animation contents.
3. Legible generated keyframes that can cover more contents in the panning shots.

Chapter 2

RELATED WORKS AND THEORIES

In general, the video summarization is a field of research dealing with shots and frames. A video sometimes contains too much irrelevant information, i.e., frames, thus video summarization can be used to extract only important information in order to save time and storage spaces. This is mostly accomplished by determining the importance of each frame, and any unimportant frames are discarded, leaving only important frames that are defined as keyframes.

2.1 Existing Approaches

Latest approaches on video summarization field concentrate on accuracy or importance of elements represented in each keyframe. Some approaches proposed a color histogram difference that separates the cut [1, 2]. It comes with the problem of flashing frames generating false cut which has to be solved with false detection using frames redundancy in a certain interval. Another approach classifies video segments into several types and employs specific treatment suitable for each type and uses entropy difference instead of color histogram [3]. These approaches manage to retrieve good keyframes under their definitions of importance. Other two approaches then improved the efficiency by reducing number of keyframes [4], and bringing audio into account as another feature [5]. There was also an effort to improve the algorithm to be online [6].

It is obvious that all of these approaches are still restricted by the consistency of aspect ratio of keyframes. They tend to discard some information that exceeds the size of the frame or cover them by unnecessarily splitting a shot sequence into two or more keyframes. Some approaches may not strict to the size of the keyframes, but still restrict themselves to the aspect ratio [7].

2.2 Related Theories

There are some characteristics of cartoon animations in comparison with general movie videos that motivate the proposed method. First, cartoon animations are not real. Some previous algorithms dealing with luminance, perspectives, dimensions, or view angles in the movies cannot be applied to every cartoon animation. Second, cartoon animations, unlike the movies, usually do not contain a long shot sequence.

The shot boundary detection is another approach that can be applied to separate a video into shot sequences based on transition of shots. Le, et al. proposed a method [8] to detect shot boundary and the results satisfied general cases of shot transitions. Their method is used to detect shot boundary for general cases in the proposed method. In other cases, shot sequences can be separated by using optical flow which is a field of vectors represents motions of objects in two consecutive frames. In this research, the classic and well-known Lucas-Kanade method [9] is implemented to retrieve the optical flow of sample videos.

A panorama keyframe, as a solution proposed in this research, is a technology [10] that stitches two or more panning frames together using some overlapping elements on those frames. The result of panorama stitching is a frame that contains contents from all frames involved in the stitching. It is expectedly wider or taller than the input videos' aspect ratio depending on the direction of panning. The problem with continuous information as mentioned earlier can be solved by these generated keyframes.

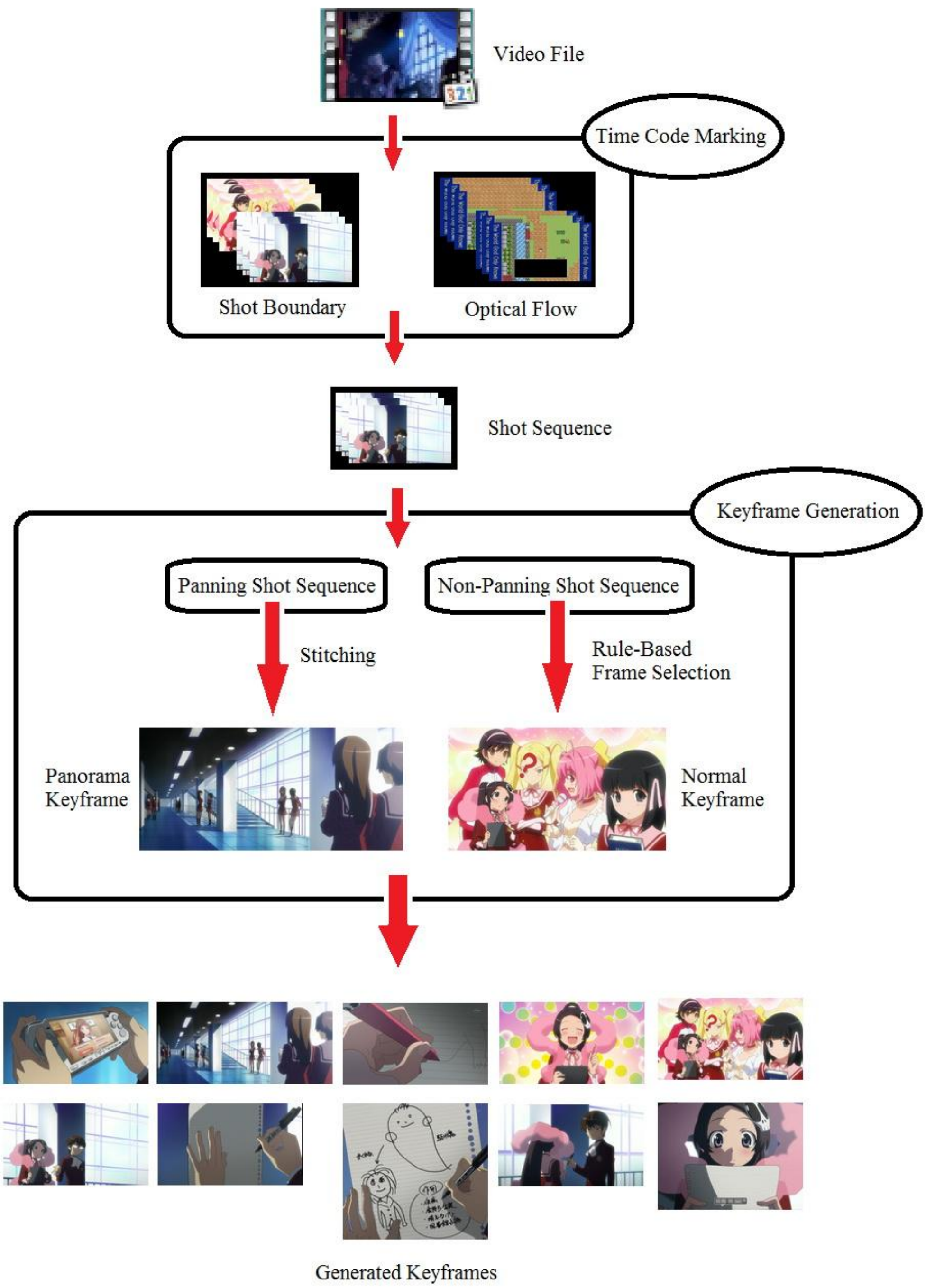

Figure 2.1: The proposed framework. ©Tamiki Wakaki/Shogakukan, Kaminomi project, TV TOKYO

Figure 2.1 shows the proposed framework for keyframe generation. The video file is passed through the Time Code Marking process to obtain shot sequences under shot boundary and optical flow techniques. The shot sequences will then be classified as “panning” or “non-panning”. The Keyframe Generation step will then generate keyframes according to shot sequences’ classes. These keyframes will be organized into comic pages later.

Parts of this thesis were presented and published in academic conferences at IEEE-International Symposium on Intelligent Signal Processing and Communication Systems (IEEE-ISPACS) 2011, Chiangmai [11] and IS&T/SPIE Electronic Imaging 2012, Burlingame, California [12].

Chapter 3

PROPOSED METHOD

This section describes how keyframes can be generated. First, the input video will have its time code marked from time to time using two different techniques: shot boundary and optical flow as described in section 3.1. A collection of frames between two consecutive marked time codes will be referred to as a "shot sequence" from now on. Keyframes are then generated from these shot sequences using rule-based algorithm as proposed in section 3.2 which is designed to optimize the keyframe generation based on cartoon animation characteristics. These keyframes will then be organized as proposed in section 3.3.

3.1 Time Code Marking

In this step, the time is marked to form a time interval which is used to indicate a shot sequence. There are two assumptions used to mark a time interval of a shot sequence, thus two techniques are used in the proposed method: shot boundary and optical flow. Shot boundary is used to indicate the end/start of a shot sequence when two consecutive frames contain distinct contents, whilst optical flow is used to indicate a shot sequence from camera panning action. Examples of time code marking are shown in Figure 3.1 and 3.2.

$$\mu = \frac{1}{N}\sum_{P=1}^{N} C_P \tag{3.1}$$

$$\sigma = \left[\frac{1}{N}\sum_{P=1}^{N}(C_P - \mu)^2\right]^{\frac{1}{2}} \tag{3.2}$$

$$s = \left[\frac{1}{N}\sum_{P=1}^{N}(C_P - \mu)^3\right]^{\frac{1}{3}} \tag{3.3}$$

3.1.1 Shot Boundary

Le, et al. [8] deploys two features to detect the shot transition; such as Color Moments (GCM) and Edge Orientation Histogram (EOH).

The GCM in equations 3.1, 3.2, and 3.3 represents the color distribution of a certain frame based on mean (μ), variance (σ), and skewness (s) of the color value of a pixel compared to every pixel as shown in the following equations where C_P is the value of the color component of P^{th} pixel, and N is the number of pixels in an image.

The EOH is a histogram of frame's edge detected by Canny edge detector. It is suggested that edge direction in a frame is interesting, thus the reason EOH is deployed. With these two features, the distances between each frame and its neighboring frames are calculated.

Using shot boundary, the time code is marked based on shot transition. The time is marked when the shot transition occurs as shown in Figure 3.1. Because shot transition represents a point where the contents in two consecutive frames abruptly change, the panorama generated from stitching these frames together will give some undesired garbled keyframe. However, shot boundary cannot mark the time code if there is no shot transition as in the case when a camera pans, especially when the panning changes the direction. In this case, optical flow can be used to solve the problem.

Figure 3.1: Examples of time code marking at shot transition. ©Tamiki Wakaki/Shogakukan, Kaminomi project, TV TOKYO

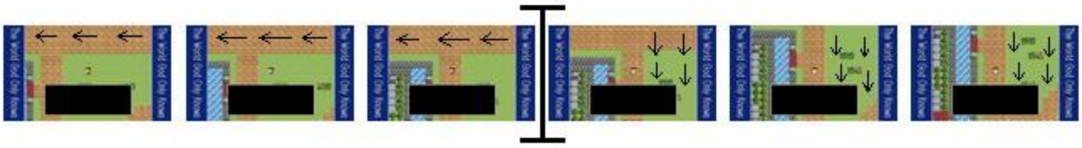

Figure 3.2: Examples of time code marking at change of optical flow. ©Tamiki Wakaki/Shogakukan, Kaminomi project, TV TOKYO

3.1.2 Optical Flow

Based on Lucas-Kanade's representation [9], the optical flow is represented as a matrix of vectors as shown in equation 3.4 where V_{YX} indicates the vector in the YX^{th} pixel.

$$\begin{bmatrix} V_{11} & \cdots & V_{1X} \\ \vdots & \ddots & \vdots \\ V_{Y1} & \cdots & V_{YX} \end{bmatrix} \tag{3.4}$$

The vector $V_{YX} = V_{180,320}$ (according to video aspect ratio of 320 x 180) of each pixel indicates the motion of objects or edges contained in the animated image. In the simplest explanation, each vector is the distance and the direction of an object or an edge travel by a specific number of frames.

$$\overline{V_C} = (X_C Y_C, X_{C-F} Y_{C-F}) \tag{3.5}$$

The above equation indicates a vector in the current frame V_C equals to the vector from current YX^{th} pixel to the previous pixel where F denotes the number of frames between the current one and the previous one.

The time code is marked based on global optical flow direction. While the algorithm uses the shot boundary to mark the time code, it also uses the optical flow to mark the time code in case there is no sudden change in the content but a shot is panning in multi-direction. In general, if a shot is panning, most of the vectors in the optical flow (also called “global optical flow”) should point to the same direction with the same magnitude and the direction of a shot sequence is always in the opposite direction of the global optical flow. Optical flow that represents vectors in other directions is called local optical flow. If the panning changes direction, the global optical flow will change its direction as well. In this case, each shot sequence should contain a shot panning in just one direction thus the global optical flow can be used to mark the time code whenever the global optical flow changes the direction. Figure 3.2 shows the frame sequence with black arrows indicate the panning direction. The sequence is marked when the shot is panning to the left, then suddenly changes the direction downward; therefore, the shot sequence should be divided into two shot sequences.

Figure 3.3, 3.4, 3.5, 3.6, and 3.7 shows the visualization of optical flow where the arrows illustrate the motion of objects which always move in the opposite direction of a camera motion. The length of arrows indicates the magnitude of motion vector, in other words, the velocity of object motion. Figure 3.3 illustrates the optical flow where the global optical flow points upward and the local optical flow (circled) points to the left indicates hair swaying. Figure 3.4 illustrates the global optical flow points upward while the camera is panning down without any local optical straying off its way. Figure 3.5 and 3.6 illustrate the global optical flow points to the left while the camera is panning to the right with all the vectors in global optical flow point to the same direction with the same magnitude. Figure 3.7 illustrates the optical flow points out to the boundary of the frame while the camera is zooming in.

Figure 3.3: Optical flow visualization where the camera is panning downward with some local optical flow circled. ©Tamiki Wakaki/Shogakukan, Kaminomi project, TV TOKYO

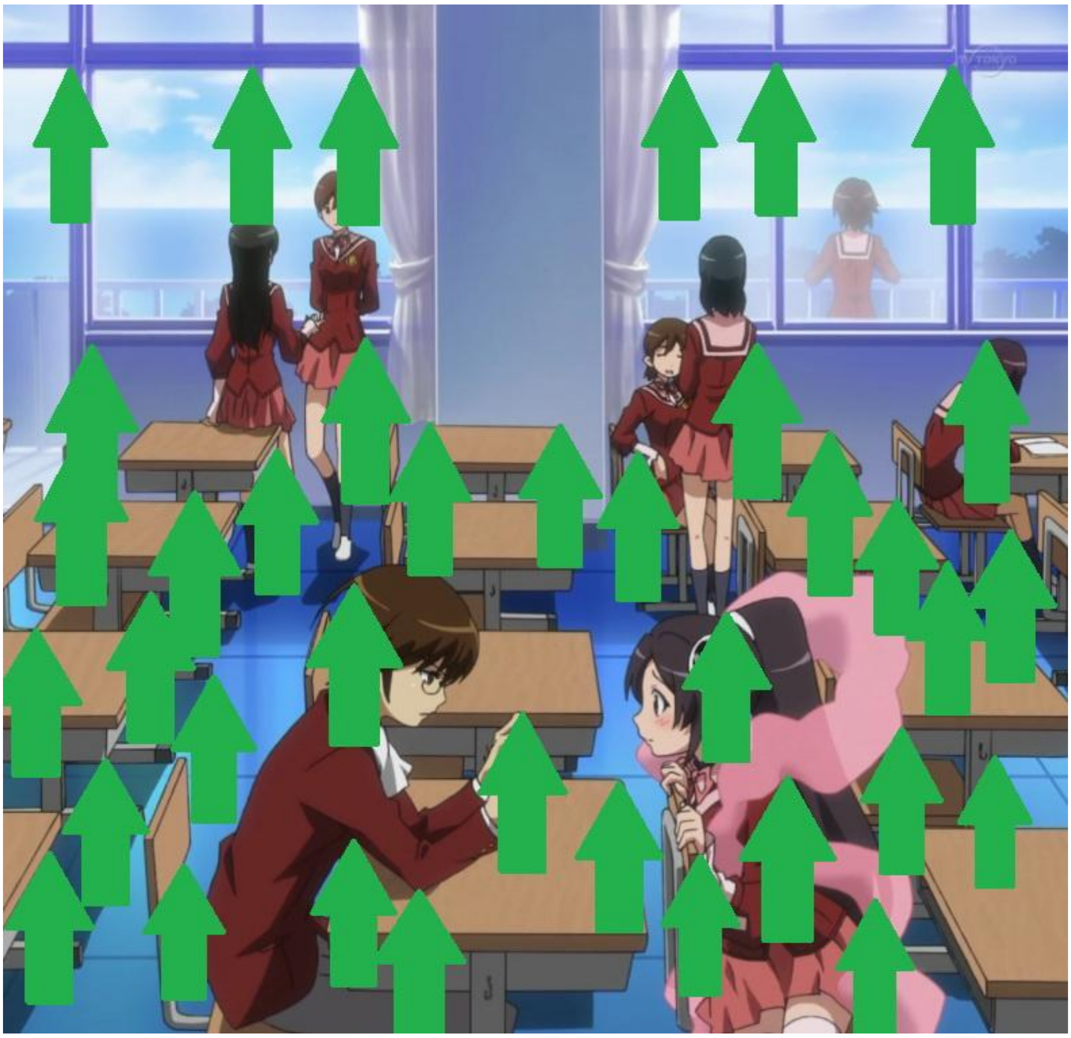

Figure 3.4: Optical flow visualization where the camera is panning downward without any straying local optical flow. ©Tamiki Wakaki/Shogakukan, Kaminomi project, TV TOKYO

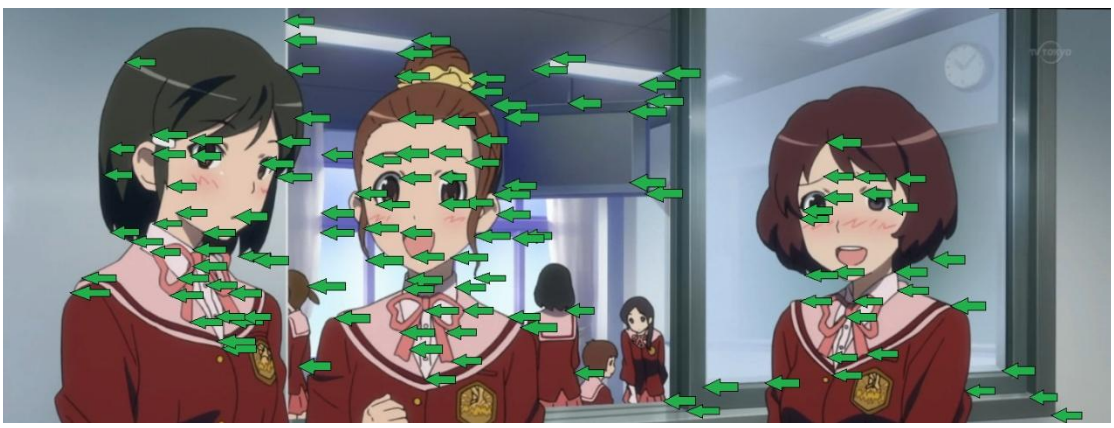

Figure 3.5: Optical flow visualization where the camera is panning to the right. ©Tamiki Wakaki/Shogakukan, Kaminomi project, TV TOKYO

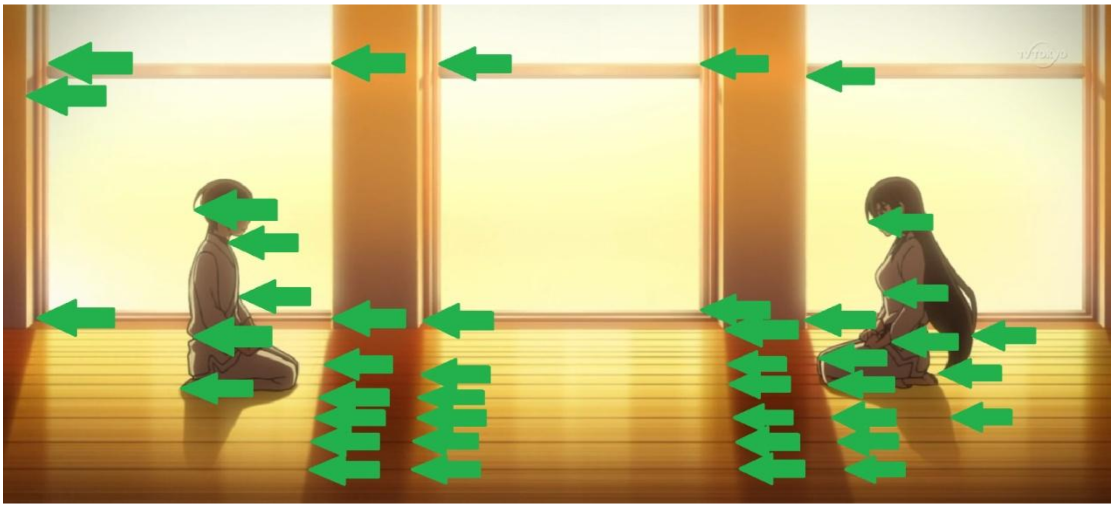

Figure 3.6: Optical flow visualization where the camera is panning to the right. ©Tamiki Wakaki/Shogakukan, Kaminomi project, TV TOKYO

Figure 3.7: Optical flow visualization where the camera is zooming in. ©Tamiki Wakaki/Shogakukan, Kaminomi project, TV TOKYO

Figure 3.8: Example keyframes obtained from SCC shot. ©Tamiki Wakaki/Shogakukan, Kaminomi project, TV TOKYO

3.1.3 False Reduction

It is common that cartoon animation usually contains some visual effects that cause redundancy on detected keyframes known as Short Color Changes (SCC). This occurs in some certain circumstances that the single shot is broken into several shots by some effects that cause drastic change in color feature; e.g. explosions, flashes, negatives, alternate images. Sample keyframes obtained from SCC shot are shown in Figure 3.8. These five keyframes are obtained by shot boundary detection in the shot that uses two images alternatively. This is supposed to be optimized into only two keyframes.

The SCC detection algorithm is employed after the time code is marked in order to reduce the redundant keyframes. It compares the current frame F_C to the next L frames F_{C+1} to F_{C+L} which are the farthest possible neighbor frames for SCC according to the experiment Ionescu et al. commenced [2]. This reduces the number of redundant frames down to no more than two frames. The proposed method once again uses the same features as used in section 3.1.1 to compare the similarity between the frames in question.

Figure 3.9: An example of dynamic shot that the global optical flow cannot be determined. ©Tamiki Wakaki/Shogakukan, Kaminomi project, TV TOKYO

3.2 Keyframe Generating

In this step, keyframes are generated from shot sequences obtained in the previous section using rule-based optical flow. Each shot sequence should contain global optical flow which is a field of the majority vectors that point to the same direction with the same magnitude. This global optical flow usually represents the scene or background since it is the major component of a frame. Optical flows in other directions are called local optical flows and they usually represent movements of objects. By considering shot sequences using the optical flow, shot sequences can be classified into two categories: panning and non-panning sequences. If most of the vectors in a shot sequence point to the same direction with the same magnitude as global optical flow then a sequence is said to be “panning” else if a sequence contains global optical flows that stay still for a period of time then it is said to be a “non-panning” sequence. There is also a case when optical flows in a shot sequence point to various directions, this shot sequence is called “dynamic”. A dynamic shot sequence as shown in Figure 3.9 is assumed to contain local optical flows that represent the movements of objects in a shot sequence; thus, a shot sequence is classified as “non-panning”.

3.2.1 Non-panning

When a shot sequence is not panning, only one frame is selected as a keyframe according to the rule-based optical flow as described below. Since optical flow provides information about the movements of objects in a shot sequence, this information can be used as keyframe selection criteria. Thus there are four rules for keyframe selection of a non-panning shot sequence.

1. Dynamic Shot Sequence: In dynamic shot sequence, there are lots of objects moving around in a shot sequence that create multi-direction optical flows. This shot sequence usually contains frames with minor difference in the content. The main content of this shot sequence usually contains objects moving in static background. Thus it is optimal to select only one frame from a shot sequence because most of the frames are redundant. The object motions usually span throughout the shot sequence, for example, moving from one side to the other side of a frame. Therefore, it is common to select the middle frame to represent a moving object.

2. Static Shot Sequence with Object Moving Out: For this type of shot sequence, global optical flow stays still but the shot sequence may contain some local optical flows which represent objects moving out of the frame. The frame contains the first motion is selected as a keyframe to depict the content because objects might be out of range in other frames.

3. Static Shot Sequence with Object Moving In: In contrast to the previous case, this type of shot sequence occurs when local optical flows show that objects are moving into the frame. For this case, the frame contains the last motion is selected as a keyframe to depict the last position of objects which is the most important content of a shot sequence.

4. Zooming Shot: This type of shot sequence occurs when most of optical flows are pointing into or pointing out from some specific location, depict objects shrinking or enlarging. In this case, the frame that covers most contents is selected as a keyframe. In other words, the first frame will be selected when the shot is zooming in, and the last frame will be selected when the shot is zooming out.

3.2.2 Panning

If the shot sequence is panning, a panorama will be generated. In order to generate a panorama keyframe, at least two frames must be selected from a shot sequence. Intuitively, the first and the last frames from a shot sequence are selected with additional frames if necessary. Once the first frame is selected, the algorithm determines the magnitude of a vector between two consecutive frames starting from the first keyframe if the magnitude is greater than half of the frame width or height [$|V| > 320/2$ or $180/2$ horizontally or vertically] then the algorithm will select the latter frame as an additional frame. The algorithm then uses the same criteria to recursively determine the magnitude of a vector between this additional keyframe and the next frame until the last frame of a shot sequence is reached. The last frame is always selected as the last keyframe. These selected keyframes are then stitched together using the existing technique [10]. There are some special cases to be treated individually:

1. For panning when movements of objects occur before or after the panning in the same shot sequence thus creating local optical flows, i.e., a panning shot sequence with local optical flow, these movements may hold important contents in some cases and cannot be neglected. In this case, the sub-sequences with movement are treated as separated non-panning sequences and fall into the non-panning category.

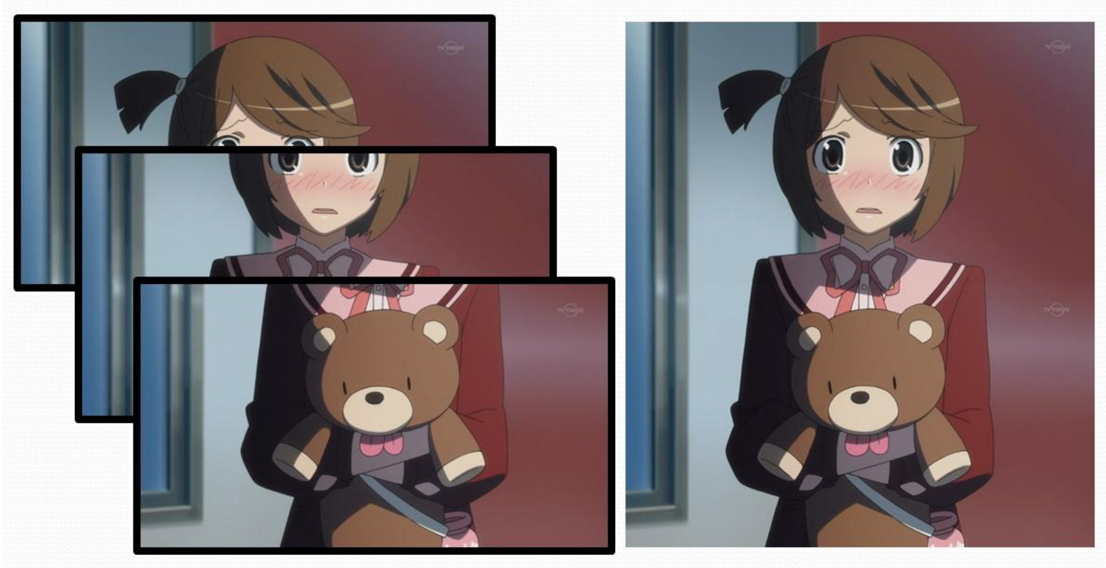

Figure 3.10: The selected consecutive keyframes versus the stitched panorama keyframe. ©Tamiki Wakaki/Shogakukan, Kaminomi project, TV TOKYO

2. When local optical flows are detected while the camera is panning or the panning shot is moving non-linearly, only the first frame and the last frame will be kept. This is to avoid the distortion as described in section 4.1.

Figure 3.10 shows how the keyframes are selected and stitched. Stitching this example, for an instant, needs the middle keyframe to provide some information in the middle of panning. Only the first and the last keyframes selected would be inadequate to track the panning direction, thus the panorama stitching cannot be carried out.

3.3 Panel Organizing

After the keyframes are generated from section 3.2, they will be organized in a comic style as shown in Figure 3.11 in such a way that each obtained keyframe becomes a comic panel. Panel organizing is restricted to no more than two columns and four rows for a clear depiction; otherwise, a respective panel will be too small. A comic page is initialized to a standard dimension 2:3. There are three types of arrangements as shown in Figure 3.11, and there are five types of classified panels.

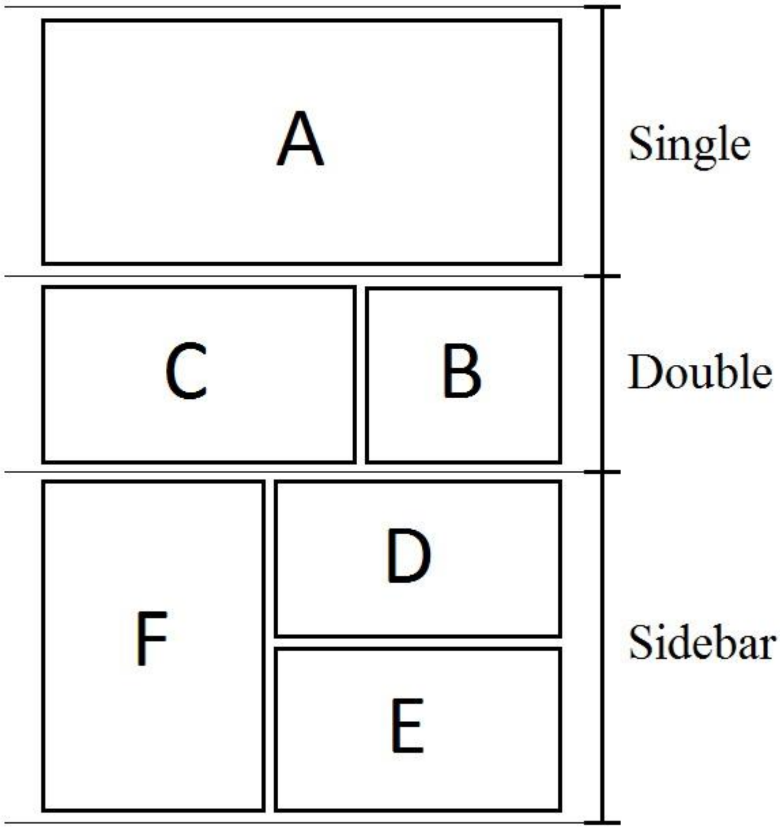

Figure 3.11: Types of arrangements

1. Normal Panel: obtained from non-panning keyframe. Panel C, D, and E in Figure 3.11 are examples of Normal Panel.

2. Wide Panel: obtained from horizontal panning keyframe. Panel A in Figure 3.11 is an example of Wide Panel.

3. Semi-Wide Panel: obtained from vertical panning keyframe with dimension wider than 11:9.

4. Square Panel: obtained from vertical panning keyframe with dimension narrower than 11:9 but not narrower than a square. Panel B in Figure 3.11 is an example of Square Panel.

5. Tall Panel: obtained from vertical panning keyframe with dimension narrower than a square. Panel F in Figure 3.11 is an example of Tall Panel.

```
--------------------------------------------------------
Algorithm 1. Panel Organizing
--------------------------------------------------------
panelOrganizing(P)
while i < n
      if(isWide(Pi))||(isSemiwide(Pi))
            setSingle(Pi)
            i=i+1
      else if(isNormal(Pi)&& isNormal(Pi+1))
            if(isNormal(Pi+2)&& isNormal(Pi+3))
                  setDouble(Pi,Pi+1)
                  setSingle(Pi+2)
                  i=i+3
            else if(isTall(Pi+2))
                  setLeftsidebar(Pi+2)
                  i=i+3
            else if(isSquare(Pi+2)||isSquare (Pi+3))
                  setDouble(Pi,Pi+1)
                  setDouble(Pi+2,Pi+3)
                  i=i+4
            else
                  setSingle(Pi,Pi+1,Pi+2)
                  i=i+3
      else if(isNormal(Pi)
                  &&(isSquare(Pi+1)||isTall(Pi+1)))
            setDouble(Pi,Pi+1)
            i=i+2
      else if(isSquare(Pi))
```

```
        setDouble(Pi,Pi+1)
        i=i+2
    else if(isTall(Pi))
        if(isNormal(Pi+1)&& isNormal(Pi+2))
            setRightsidebar(Pi)
            i=i+3
        else if (isSquare(Pi+1)||isTall(Pi+1))
            setDouble(Pi,Pi+1)
            i=i+2
        else
            setSingle(Pi,Pi+1,Pi+2)
            i=i+3
else
    setSingle(Pi,Pi+1,Pi+2)
    i=i+3
---------------------------------------------------------
```

Initially, it is regulated that a page contains four rows for optimization, except the cases where a page contains only one or no double at all. When a panel is single, it consumes more spaces in the page, thus the number of rows in a page is reduced to three in order to fit the spaces.

With the definitions and initial settings mentioned earlier, the panel organizing is executed as described in Algorithm 1, where P indicates all panels from generated keyframes, and P_i is the i^{th} panel and n is the number of all panels obtained from the keyframes in previous sections. The algorithm is executed starting from the first panel P_1 in a collection and the value i is incremented along the process. This algorithm is designed to organize the panels using general comics as an example, whereas it tries to optimize the page spaces as much as possible. It also orders the panels so as to reduce confusion to the very least.

Figure 3.12: An example of organized comic page. ©Tamiki Wakaki/Shogakukan, Kaminomi project, TV TOKYO

Figure 3.12 shows an example of the finished comic page that consists of Normal and Tall panel in the top row, two Normal panels in the second row, a Normal panel in the third row, and a Wide panel in the bottom row.

Chapter 4

EXPERIMENTAL RESULTS AND DISCUSSION

The accuracy and performances of the proposed method in video summarization aspects are evaluated on various types of cartoon animation videos. There are four major types of videos: high-quality action, low-quality action, high-quality non-action, and low-quality non-action. These four types hold different characteristics. The actions usually consist of more shot transitions and mostly dynamic shots compared to non-action. The quality of the video determines how elaborate the frames are depicted. Low-quality video usually comes with distorted or blurred in-between frames that maybe selected along the process. On the other hand, the high-quality one usually comes with detailed motions that can affect the results.

The sample videos are decoded into MPEG videos that contain only chunks of frames without audio part, and passed through the process as described in chapter 3. Four types of videos are tested as samples. Each sample is 20 minutes long and consists of 28,800 frames. There are two major checkpoints regarding the proposed method's performances. The results among those types of videos are discussed below according to its checkpoints.

4.1 Panorama Keyframe

This section evaluates how efficient the proposed method can generate keyframes. The panning shot sequences are shown to be accurately and perfectly converted into panorama keyframes, with some inevitable exceptions. According to Figure 4.1, the proposed method passes a frame sequence as seen in Figure 4.1(a) into the process and generates a panorama keyframe that contains all important contents from the original source as seen in Figure 4.1(b). Moreover, it can be seen that a keyframe is also stitched naturally.

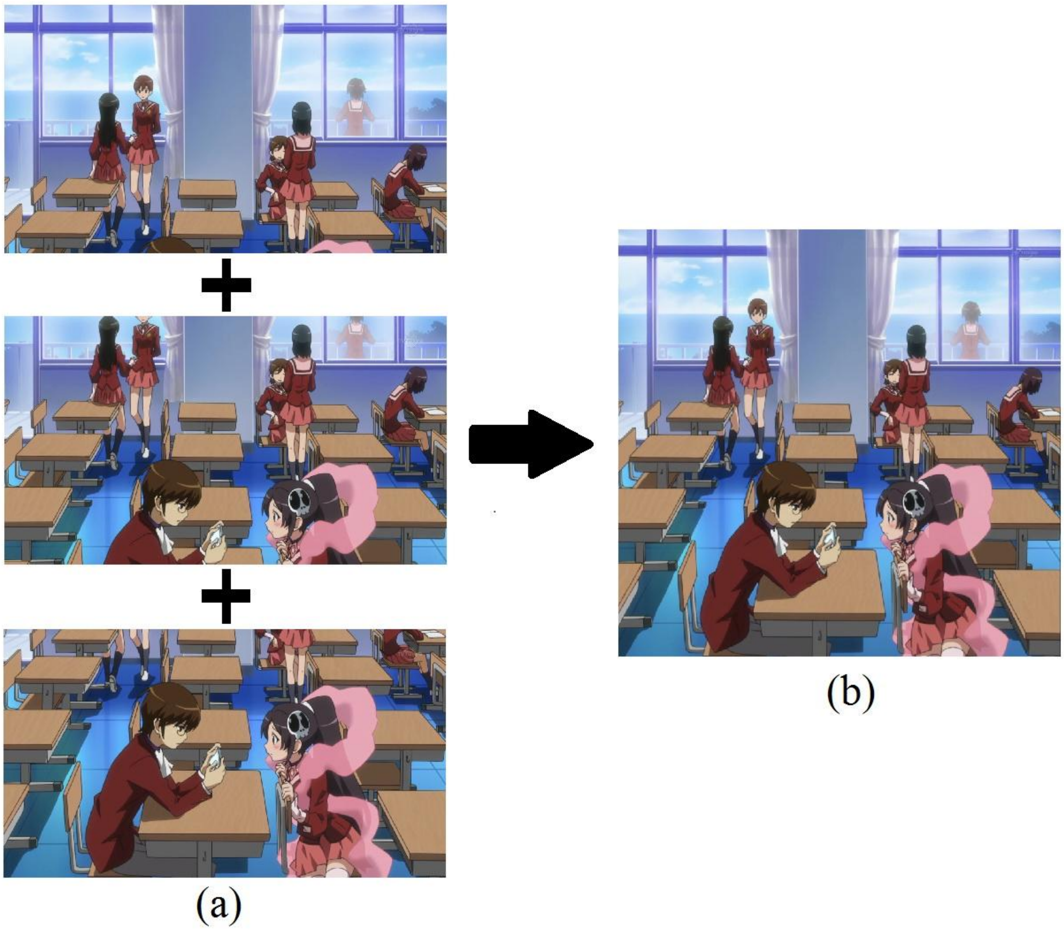

Figure 4.1: Comparison between (a) a shot sequence and (b) a stitched panorama keyframe. ©Tamiki Wakaki/Shogakukan, Kaminomi project, TV TOKYO

The proposed panorama keyframe generation process gives almost a perfect result when a shot sequence pans with minimal amount of local optical flows; i.e. objects are moving while the camera is panning. When some objects are moving along the panning scene, the frames are stitched together to generate a panorama keyframe that may contain extra elements.

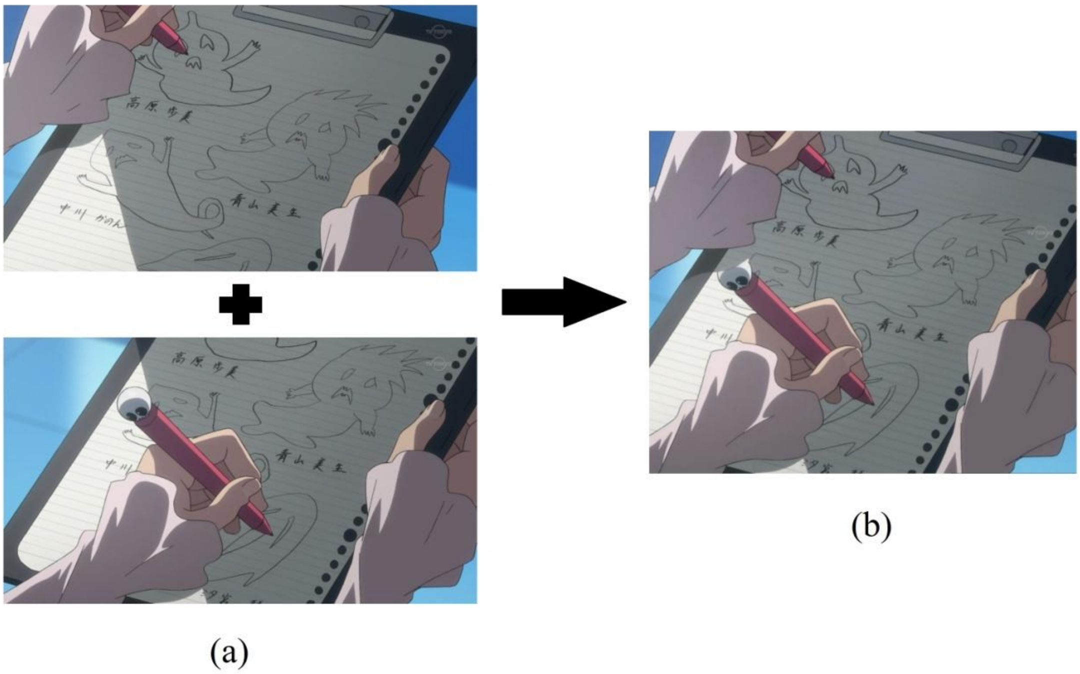

Figure 4.2: An example of a distorted panorama keyframe. (a) The left hand is moving downward from top to bottom of the frame (b) Panorama keyframe generated from two keyframes that contain the hands. ©Tamiki Wakaki/Shogakukan, Kaminomi project, TV TOKYO

Figure 4.2 shows that the algorithm detects the panning of a shot sequence and captures the first and the last frames from the shot sequence. Notice that the left hand, which is an object in this shot sequence, is moving from top of the start frame to the bottom of the end frame while the shot sequence is panning downward. When the selected frames are stitched together, the hand was visible in two locations and this panorama is classified as a distorted keyframe.

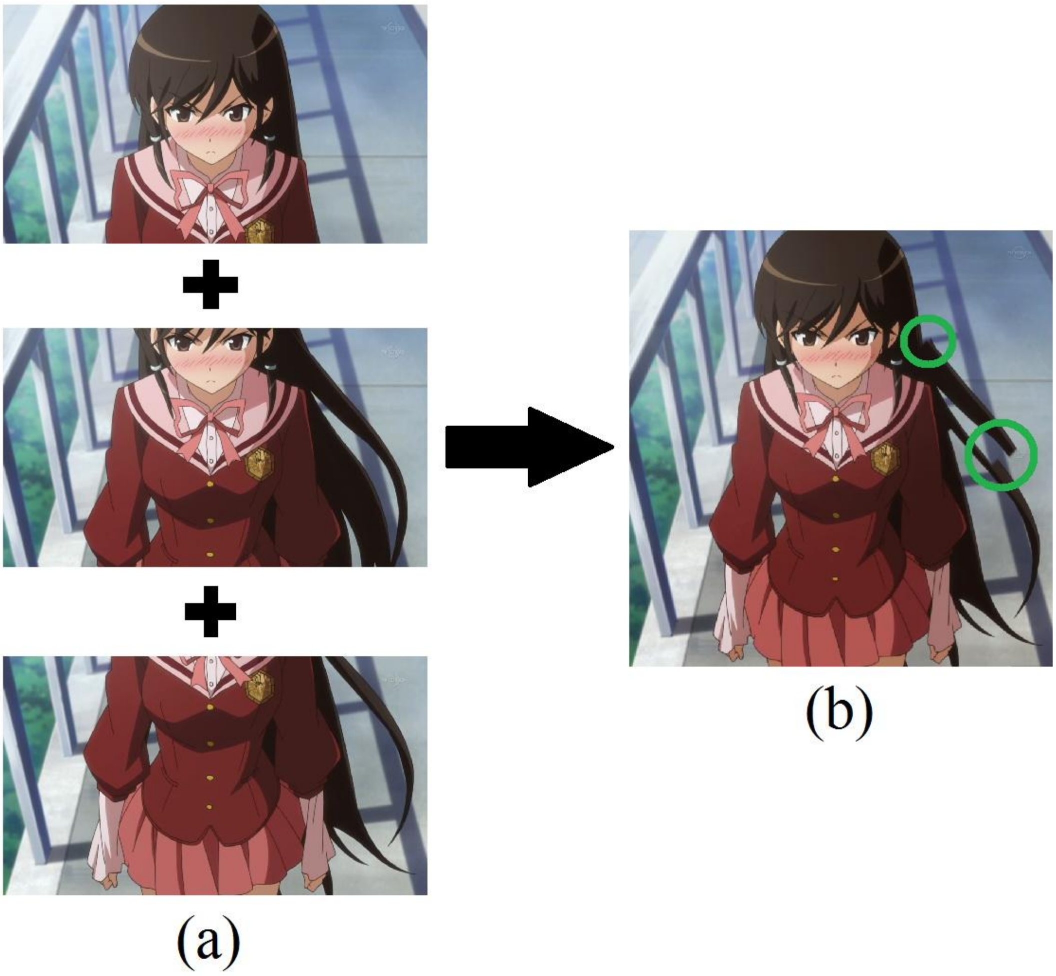

Figure 4.3: An example of a distorted panorama keyframe. (a) The hair is swaying while the camera is panning up (b) Panorama keyframe generated from two frames with distorted points circled. ©Tamiki Wakaki/Shogakukan, Kaminomi project, TV TOKYO

Figure 4.3 shows that the keyframe has some distortions on the girl's hair. This is because the algorithm detects a panning shot sequence and captures the first and the last frames from the shot sequence, while the girl's hair is swaying making the algorithm stitched them incompletely. This panorama keyframe is also classified as a distorted keyframe.

Figure 4.4: An example of a keyframe sequence when the camera is panning around an object, ordering from right to left of the top row then right to left of the bottom row. ©Tamiki Wakaki/Shogakukan, Kaminomi project, TV TOKYO

Another limitation to panorama keyframe generation is when the camera panning revolves around objects. It is common that human's eyes can perceive an object in only one view, thus it is impossible for one object to be seen from back, side, and front views simultaneously. Even a model in 3D must be rotated around in order to be seen from every angle, however, it can still be seen from one angle at a time. Thus, when camera is panning around an object, a panorama keyframe cannot depict a 3D object naturally.

Figure 4.4 illustrates the sequence of keyframes obtained from the shot with the camera helically pans around the depicted girl. It is not possible to sensibly include every angle into one keyframe. It is also not viable to choose a certain angle as a representative of the shot as the camera always pans away from the stated angle, thus there is not adequate information from any angle.

Figure 4.5: An example of a keyframe generated from a wobbly panning shot sequence. ©Tamiki Wakaki/Shogakukan, Kaminomi project, TV TOKYO

Table 4.1: Comparison between generated keyframes and distorted keyframes.

Video type	Keyframes	Distortions			% Good frames
		By limitation	By quality	Others	
HQ-action	693	15	12	1	95.96%
LQ-action	757	5	24	0	96.17%
HQ-non-action	496	1	1	0	99.60%
LQ-non-action	515	4	0	0	99.22%

Table 4.1 shows the experimental results. There are four types of sample videos. The table compares the number of generated keyframes to the number of distorted keyframes. Most of distorted keyframes are caused by limitations mentioned earlier together with some other distorted keyframes generated from wobbly panning shot sequences as illustrated in figure 4.5. At least more than 95% of keyframes make sense and are legible. Some selected frames from dynamic shots are blurred due to quality of animation and these are beyond the control of the proposed method.

The results also show that the non-action videos are easier for the proposed method to generate perfect keyframes. This is obviously because the action-based videos usually come with more object movements and dynamic shot sequences that cause distortion on generated keyframes.

(a) (b)

Figure 4.6: Comparison between (a) comic book version and (b) results from the proposed method after the panel organizing algorithm is implemented. ©Tamiki Wakaki/Shogakukan, Kaminomi project, TV TOKYO

Table 4.2: Generated keyframes compared with comic panels.

Video type	Matched frames	Generated keyframes		Comic panels	
		Number	%Matched	Number	%Matched
HQ-action	621	693	89.61%	672	92.41%
LQ-action	660	757	87.19%	698	94.56%
HQ-non-action	464	496	93.55%	489	94.89%
LQ-non-action	470	515	91.26%	482	97.51%

4.2 Content Accuracy

This section compares the final results from the proposed method with the official comic counterparts of the animated samples. The chosen sample videos are cartoon animations that have their own comic book adaptations with the same contents which can be used as an ideal model to evaluate the results of the proposed method. Figure 4.6 shows how the contents are exactly generated from the proposed method compared with the comic book adaptation.

The circled numbers for each panel represent order of panels in the comic book and keyframes obtained from the proposed method. This comparison shows that most of the contents generated from the proposed method are close to the comics. The order of scenes and their contents are depicted accurately according to their examples. Some frames even have their aspect ratio altered by panorama process to match their respective panels (pair no. 8). The percentages of matched frames are shown in Table 4.2.

The numbers in the Matched frames column denote the total number of panels such that their contents matched exactly with their corresponding keyframes. The numbers in the Generated keyframes column denote the total number of keyframes generated by the proposed method together with the percentages of matched frames. Likewise, the numbers in the Comic panels column denote the total number of panels featured in the comic books together with the percentages of matched frames. The proposed method manages to match at least 90% of the ideal contents in any cases.

Table 4.2 also shows that the number of keyframes generated by the proposed method is always greater than the number of comic panels which causes the degree of accuracy to be lower. This is because cartoon animations usually use several shot sequences to cover the contents while these contents are able to be crammed into one panel in comic books. This means that the percentages of matched frames when compared with the generated keyframes are less significant. It is more important to be able to cover as many contents in the ideal model as possible. The surplus keyframes aside from those contents usually add more details into it.

Chapter 5

CONCLUSIONS

This research proposes a novel method to generate keyframes from cartoon animations. Concerning the lacks of contents coverage in existing [1, 2, 3, 4, 5, 6, 7] fixed-aspect-ratio video summarization methods, this research proposes a rule-based optical flow method to generate panorama keyframes in order to increase the contents featured in each keyframe. This method is designed to especially enhance cartoon animation video summarization, as it exploits some unique characteristics that are differed from other types of video media.

The proposed method mark the time code of a cartoon animation video based on shot boundary and optical flow direction to separate periods of time into shot sequences. Each shot sequence is then passed through the keyframe generating algorithm. By using the global and local optical flow information, the algorithm classifies shot sequences and processes them according to the regulated rules, resulted in keyframes with various sizes. Each of these keyframes is then treated as comic panel and organized together into comic pages.

The results are examined under two major aspects. First, stitched panorama keyframes based on optical flow information turn out to be mostly legible and acceptable. Only minority keyframes are distorted due to the limitations of the proposed method and video quality. Second, the whole keyframe sequence is compared with its comic book adaptation and the experimental results show that the proposed method manages to generate keyframes that cover most of the contents featured in the comic books. Almost all of the keyframes exactly match the contents in the ideal model which yields a close-to-perfection achievement. These two evaluations proved that both objectives of this thesis are achieved. In addition, those generated panels are also organized into comic pages, letting users experience a representation similar to the real comic books.

REFERENCES

[1] Ionescu,B., Buzuloiu, V., Lambert, P., Coquin, D. Improved Cut Detection for the Segmentation of Animation Movies. International Conference on Acoustics, Speech, and Signal Processing, IEEE Computer Society, Honolulu, Hawaii, 2007.

[2] Ionescu,B., Lambert, P., Coquin, D., Buzuloiu, V. The Cut Detection Issue in the Animation Movie Domain. Journal of Multimedia 2(4) (2007).

[3] Mentzelopoulos, M. and Psarrou, A. Key-Frame Extraction Algorithm using Entropy Difference. International Workshop on Multimedia Information Retrieval, ACM, New York, 2004.

[4] Dumont E., and Mérialdo, B. Sequence Alignment for Redundancy Removal in Video Rushes Summarization. Proceedings of the 2nd ACM TRECVid Video Summarization Workshop, ACM, Vancouver, British Columbia, 2008.

[5] Liu, Y., Liu, Y., Ren, T., Chan, K.C.C. Rushes Video Summarization using Audio-Visual Information and Sequence Alignment. Proceedings of the 2nd ACM TRECVid Video Summarization Workshop, ACM, Vancouver, British Columbia, 2008.

[6] Abd-Almageed, W. Online, Simulteneous Shot Boundary Detection and Key Frame Extraction for Sports Videos Using Rank Tracing. 15th IEEE International Conference on Image Processing, IEEE Computer Society, San Diego, California, 2008.

[7] Boreczky, J., Girgensohn, A., Golovchinsky, G., and Uchihashi, S. An Interactive Comic Book Presentation for Exploring Video. CHI Letters (2000).

[8] Le, D.D., Satoh, S., Ngo, T.D., Duong, D.A. A Text Segmentation Based Approach to Video Shot Boundary Detection. International Workshop on Multimedia Signal Processing, IEEE Computer Society, Cairns, Queensland, 2008.

[9] Lucas, B.D. and Kanade, T. An Iterative Image Registration Technique with an Application to Stereo Vision. Proceedings of Imaging Understanding Workshop (1981) : 121-130.

[10] Ha, S.J., Koo, H.I., Lee, S.H., Cho, N.I., Kim, S.K. Panorama Mosaic Optimization for Mobile Camera Systems. IEEE Transactions on Consumer Electronics 53(4), IEEE

Computer Society, 2007.

[11] Tanapichet, P., Cooharojananone, N., Lipikorn, R. Automatic Comic Strip Generation Using Extracted Keyframes from Cartoon Animation. International Symposium on Intelligent Signal Processing and Communication Systems, IEEE, Chiangmai, 2011.

[12] Tanapichet, P., Cooharojananone, N., Lipikorn, R. Keyframe Generation from Cartoon Animation Using Rule-Based Optical Flow. Multimedia on Mobile Devices 2012; and Multimedia Content Access: Algorithms and Systems VI, IS&T/SPIE Electronic Imaging, Burlingame, California, 2012.

APPENDICES

APPENDIX A

EXPERIMENTAL SAMPLES

There are 8 samples of video for experiments. Each of them has video properties as described in chapter 4.

Table A1: Comparison between generated keyframes and distorted keyframes.

Sample	Video type	Keyframes	Distortions			% Good frames
			By limitation	By quality	Others	
FTZ05	HQ-action	340	11	6	1	94.71%
STD01	HQ-action	353	4	6	0	97.17%
ARA01	LQ-action	399	2	13	0	96.24%
TMD15	LQ-action	358	3	11	0	96.09%
HNA02	HQ-non-action	249	0	0	0	100%
TAR01	HQ-non-action	247	1	1	0	99.19%
KMN21	LQ-non-action	253	1	0	0	99.60%
SAK14	LQ-non-action	262	3	0	0	98.85%

Table A2: Generated keyframes compared with comic panels.

Sample	Video type	Matched frames	Generated keyframes		Comic panels	
			Number	%Accuracy	Number	%Accuracy
FTZ05	HQ-action	305	340	89.71%	330	92.42%
STD01	HQ-action	316	353	89.52%	342	92.40%
ARA01	LQ-action	344	399	86.22%	366	93.99%
TMD15	LQ-action	316	358	88.27%	332	95.18%
HNA02	HQ-non-action	231	249	92.77%	245	94.29%
TAR01	HQ-non-action	233	247	94.33%	244	95.49%
KMN21	LQ-non-action	212	253	83.79%	224	94.64%
SAK14	LQ-non-action	258	262	98.47%	258	100%

APPENDIX B

EXAMPLES OF VARIOUS RESULTS

Figure B1: A comic page that consists of Semi-Wide panel in the top row, two Normal panels in the middle row, and a Normal panel in the bottom row. ©Tamiki Wakaki/Shogakukan, Kaminomi project, TV TOKYO

Figure B2: A comic page that consists of two Normal panels in the top row, a Normal panel in the second row, a Normal panel with a Semi-Wide panel in the third row, and a Normal panel in the bottom row. ©Tamiki Wakaki/Shogakukan, Kaminomi project, TV TOKYO

Figure B3: A comic page that consists of two Normal panels in the top row, a Wide panel in the second row, two Normal panels in the third row, and a Normal panel in the bottom row. ©Tamiki Wakaki/Shogakukan, Kaminomi project, TV TOKYO

Figure B4: A comic page that consists of two Normal panels in the top row, a Normal panel in the middle row, and a Wide panel in the bottom row. ©Tamiki Wakaki/Shogakukan, Kaminomi project, TV TOKYO

Figure B5: A comic page that consists of all Normal panels caused by space restriction of the last page. ©Tamiki Wakaki/Shogakukan, Kaminomi project, TV TOKYO

APPENDIX C

PERMISSION DOCUMENTS

The following are copies of documents regarding usage copyrighted materials permitted by respective organizations.

Re: เรียน บก. เรื่องการขอใช้บางส่วนของหนังสือการ์ตูนเพื่อการวิจัยเชิงวิชาการ
Back to messages

To see messages related to this one, group messages by conversation.

TKO Comics - Info Add to contacts
12/29/2011
To Pakpoom Tanapichet, Issara Thienlikit
Reply

From: **TKO Comics - Info** (info@tkocomics.com)
Sent: Thursday, December 29, 2011 11:56:31 AM
To: Pakpoom Tanapichet (kaito_dash@hotmail.com)
Cc: Issara Thienlikit (issara@tkocomics.com)

เรียนคุณภาคย์ภูมิ

ขอบคุณสำหรับการติดต่อแจ้งเรื่องเข้ามาครับ

ทางเรายินดีให้คุณภาคย์ภูมินำเนื้อหาบางส่วนจากหนังสือการ์ตูนเรื่อง เซียนเกมรักขอเป็นเทพนักจีบ ไปใช้ในวิทยานิพนธ์ครับ
อย่างไรก็ตาม รบกวนช่วยให้เครดิตกับทางสำนักพิมพ์ญี่ปุ่นและผู้เขียนด้วยการใส่ข้อความประกาศลิขสิทธิ์ภาษาอังกฤษ ที่แสดงอยู่บนปกหุ้มด้านหลังและหน้าเครดิตด้วยครับ
อาจจะใส่ในหน้าอรรถาธิบาย ท้ายเล่มวิทยานิพนธ์ก็ได้ครับ

เพื่อความสมบูรณ์ของงานภาพที่จะตีพิมพ์ในวิทยานิพนธ์ ถ้าคุณภาคย์ภูมิต้องการไฟล์ดิจิทัลสำหรับภาพประกอบจากการ์ตูนเรื่องนี้ ทางเรายินดีให้ความร่วมมือส่งให้
โดยคุณภาคย์ภูมิขอให้อาจารย์ที่ปรึกษาวิทยานิพนธ์ส่งจดหมายด้วยหัวกระดาษจุฬาลงกรณ์ฯ ขอไฟล์มาที่เราครับ

ภาณุชัย ศรีนวลนัด
บรรณาธิการบริหาร

From: Pakpoom Tanapichet <kaito_dash@hotmail.com>
To: info@tkocomics.com
Sent: Thursday, December 29, 2011 11:01 AM
Subject: เรียน บก. เรื่องการขอใช้บางส่วนของหนังสือการ์ตูนเพื่อการวิจัยเชิงวิชาการ

เรียน บก. TKO Comics

เนื่องด้วยผม นาย ภาคย์ภูมิ ตาณพิเชษฐ์ ปัจจุบันเป็นนิสิตระดับปริญญาโท คณะวิทยาศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย กำลังทำงานวิจัยเกี่ยวกับอนิเมชั่นและหนังสือการ์ตูน (มังงะ) ตามระเบียบข้อบังคับของมหาวิทยาลัยที่นิสิตต้องทำวิทยานิพนธ์และเขียนงานวิจัยเชิงวิชาการเพื่อตีพิมพ์ในวารสารวิชาการนานาชาติ ซึ่งในงานวิจัยของผม ได้มีการใช้ภาพบางส่วนของหนังสือการ์ตูนเรื่อง "เซียนเกมรักขอเป็นเทพนักจีบ" หรือ "Kami nomi zo Shiru Sekai" ในการอ้างอิงเชิงวิชาการ และงานดังกล่าวก็ได้รับการตอบรับจากองค์กรวิชาการต่างประเทศเพื่อนำไปตีพิมพ์และเผยแพร่ในระดับนานาชาติ ผมจึงจะเรียนขออนุญาตจาก TKO Comics ผู้ซึ่งถือลิขสิทธิ์ตีพิมพ์การ์ตูนเรื่องดังกล่าวแต่เพียงผู้เดียวในประเทศไทย เพื่อนำบางส่วนของเนื้อหาไปเผยแพร่ในเชิงวิชาการเท่านั้น โดยมิได้หวังผลประโยชน์เชิงพาณิชย์ใดๆทั้งสิ้น

โดยส่วนที่ใคร่ขออนุญาตนำไปใช้คือ:
ภาพจากหนังสือการ์ตูนเรื่อง "เซียนเกมรักขอเป็นเทพนักจีบ" ตอนที่ 18 (หรือตอนที่ 2 ของเล่ม 3) หน้าที่ 5 และ 6 เต็มทั้งหน้า โดยตัวผมได้ทำการซื้อหนังสือฉบับลิขสิทธิ์ของ TKO Comics และสแกนหน้าดังกล่าวออกมาเป็นไฟล์ภาพ และทำการปรับแต่งตามความเหมาะสม

จึงใคร่ขออนุญาตมา ณ ที่นี้ หาก TKO Comics ต้องการเอกสารอันใดเพื่อยืนยันว่าเนื้อหาดังกล่าวจะไม่ถูกนำไปใช้เพื่อประโยชน์ในเชิงพาณิชย์ ขอให้ติดต่อที่อีเมลนี้เพื่อแจ้งให้ผมดำเนินการได้ทันที

ด้วยความเคารพ
ภาคย์ภูมิ ตาณพิเชษฐ์
นิสิตระดับปริญญาโท คณะวิทยาศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย

ที่ ศธ 16 /2555

ภาควิชาคณิตศาสตร์และวิทยาการคอมพิวเตอร์
คณะวิทยาศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย
ถนนพญาไท ปทุมวัน กรุงเทพฯ 10330

5 มกราคม 2555

เรื่อง ขออนุญาตนำภาพบางส่วนของการ์ตูนเพื่อไปประกอบการทำวิทยานิพนธ์

เรียน กรรมการผู้จัดการ บริษัท ทีไอจีเอ จำกัด

ด้วย นายภาคย์ภูมิ ตาณพิเชษฐ์ นิสิตระดับปริญญามหาบัณฑิต สาขาวิชาวิทยาการคอมพิวเตอร์และเทคโนโลยีสารสนเทศ (หลักสูตรภาษาอังกฤษ) ภาควิชาคณิตศาสตร์และวิทยาการคอมพิวเตอร์ คณะวิทยาศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย กำลังทำงานวิจัยเกี่ยวกับภาพเคลื่อนไหว (Animation) โดยมีดิฉัน ผู้ช่วยศาสตราจารย์ ดร.รัชลิดา ลิปิกรณ์ เป็นอาจารย์ที่ปรึกษาวิทยานิพนธ์ ตามระเบียบข้อบังคับของมหาวิทยาลัยที่นิสิตต้องทำวิทยานิพนธ์และเขียนงานวิจัยเชิงวิชาการ เพื่อตีพิมพ์ในวารสารวิชาการนานาชาติ ซึ่งในงานวิจัยดังกล่าวนี้ ได้มีการใช้ภาพบางส่วนของการ์ตูนเรื่อง "The World God Only Knows Season 2" ในการอ้างอิงเชิงวิชาการ และงานดังกล่าวก็ได้รับการตอบรับจากองค์กรวิชาการเพื่อนำไปตีพิมพ์และเผยแพร่ในระดับนานาชาติ

ทั้งนี้ นายภาคย์ภูมิ ตาณพิเชษฐ์ นิสิตภายใต้ความดูแลของดิฉัน จึงใคร่เรียนขออนุญาตจากบริษัท ทีไอจีเอ จำกัด ผู้ซึ่งถือลิขสิทธิ์ในการผลิตและจัดจำหน่ายการ์ตูนเรื่องดังกล่าวแต่เพียงผู้เดียวในประเทศไทย เพื่อนำภาพบางส่วนของเนื้อหาไปเผยแพร่ในเชิงวิชาการเท่านั้น โดยมิได้หวังผลประโยชน์เชิงพาณิชย์แต่อย่างใดทั้งสิ้น โดยส่วนที่ใคร่ขออนุญาตนำไปใช้นั้น นายภาคย์ภูมิ ตาณพิเชษฐ์ ได้ทำการอัดจากโทรทัศน์ช่อง TV Tokyo ในขณะที่การ์ตูนเรื่องดังกล่าวได้ทำการออกอากาศที่ประเทศญี่ปุ่นโดยตรง

จึงเรียนมาเพื่อโปรดให้ความอนุเคราะห์ด้วย จักเป็นพระคุณยิ่ง

ขอแสดงความนับถือ

รัชลิดา ลิปิกรณ์

(ผู้ช่วยศาสตราจารย์ ดร.รัชลิดา ลิปิกรณ์)
อาจารย์ที่ปรึกษาวิทยานิพนธ์

ฉิทธิชัย รุจิภาสกุล
กรรมการผู้จัดการ 6/1/55

ภาควิชาคณิตศาสตร์และวิทยาการคอมพิวเตอร์
โทรศัพท์ 02-2185141-2 โทรสาร 02-2552287

(087) 093-5232 ห. ภาคภูมิ (นิสิตฯ)

BIOGRAPHY

Pakpoom Tanapichet graduated a bachelor degree in Electrical Engineering from Thammasat University, Thailand, in 2009 and a master degree in Information Technology and Computer Science at Chulalongkorn University, Thailand, in 2012. The concentrated fields of researches include Image Processing, Video Summarization, and Multimedia Content Recognition.